Columbia University

Department of Economics
Discussion Paper Series

Analogical Reasoning in Decision Processes

Massimiliano Amarante

Discussion Paper No.: 0506-17

Department of Economics
Columbia University
New York, NY 10027

February 2006

Analogical Reasoning in Decision Processes

Massimiliano Amarante*
Columbia University, NY†

February 5, 2006

Abstract

We give a definition of reasoning by analogy, which is tailored to a setting of decision making under uncertainty. We present a model of decision making which is based on such a definition, and show that it is compatible with a large class of preferences.

1 Reasoning by Analogy

Analogy is the recognition that one thing A (a phenomenon, a problem, etc.) is like another thing B and that, therefore, consequences (inferences, explanations, solutions, etc.) that can be drawn from A can be drawn from B as well. Analogy is a reasoning process that is so pervasive in human life that can be viewed as one of the cornerstones of human thought. In the Dictionary of Philosophy of Mind [7], Analogy is recognized as, "... an important kind of thinking contributing to such cognitive tasks as explanation, planning and decision making (Paul Thagard)". Because of this, it is not surprising that the concept of Analogy has been an object of philosophical reflection since ancient times.[1]

As of today, research on the concept of Analogy and on the process of analogical reasoning has become increasingly important in the study of Artificial Intelligence (see, for instance, [10]). Again, this is not surprising given the nature of the subject. What is surprising is that classical theories of decision making under uncertainty do not explicitly recognize the role played by analogical reasoning.

Theories of decision making under uncertainty are concerned with explaining/guiding the behavior of individuals who have to choose a course of action

*I benefited from comments and suggestions from Paolo Ghirardato, Tzachi Gilboa, Ket Richter, Aldo Rustichini, Paolo Siconolfi, Max Stinchcombe and Stan Wellisz. I also wish to thank seminar participants at RUD 2003, CalTech, University of Minnesota, University of Texas-Austin, Columbia University, University of Rochester, Rutgers University, University of Turin, Ohio State University.

†Department of Economics, Columbia University, 1019 IAB, 420 W 118th street, NY, NY 10027. Email: ma734@columbia.edu.

[1] To give just one example, reflections on the concept recur several times in St. Thomas from the Summa Theologica to the De potentia to the De veritates.

in an uncertain environment. It is indisputable that, in a large variety of situations, individuals perceive this as a "new" problem, in the sense that they do not have a "ready-made" solution for it. In such situations, individuals tend to rely on their past experience, that is, they recall solutions that they gave in the past and try to "adapt" them to the new circumstances. It is precisely these types of processes that are unaccounted for by classical theories.

In recent years, Gilboa and Schmeidler started a research program (see, for instance, the comprehensive [16] and the more recent [17]) aiming to fill, at least in part, this gap. Their Case-Based Theory of Decisions can be viewed as a specialization to a decision theoretic setting of the idea of Case-Based Reasoning. This is founded on the intuition that the problems one faces are often similar to problems encountered in the past and, therefore, that past solutions may be of use in current situations. The idea of Case-Based Reasoning becomes especially powerful when the reasoning method is coupled with a learning paradigm: solutions given to new problems update the case base thus producing new strategies to attack new problems. Strictly speaking, Case-Based Reasoning refers to problems based on single-domain cases.[2] As such, it is a special case of Analogical Reasoning that allows for the use of past cases from different domains and goes even beyond that.

In the next section, I propose a definition of Reasoning by Analogy that is tailored to a setting of decision making under uncertainty. Building on this definition, I present (Section 5) a model of decision making which accounts for the possibility that a decision maker might choose a course of action based on past or even on only potential cases from different domains. The model contains several ingredients that combine to give rise to a certain structure. In Section 6, I show that one of the most general axiomatic models of decision making we know of (see [11]) displays the same structure as the model built on the idea of analogical reasoning. In a way, that structure had always been there, and the task accomplished by Theorem 2 (Section 5) is that of bringing that structure to light. Later, in Section 7, I show that the result holds for axiomatizations weaker than that in [11]. A few comments about the result are given in Section 8. There, I also briefly discuss the link between the concept of analogical reasoning presented in this paper and some important problems in decision theory like unforeseeen contingencies and ambiguity.

2 Models and Analogies

Suppose that a decision maker has to rank a set of bets, $\mathcal{F}$, regarding the meteorological conditions in Siberia (S) over the month of July in a certain year. Lacking any direct experience of the weather in Siberia, our decision maker might try to "translate" this problem into a problem he is more familiar with. For instance, the meteorological conditions over the same month in his own state, which, we suppose, is California (C). In other words, let us assume

[2] Roughly speaking, different cases pertain to a single domain if they are describable by means of the same set of parameters.

that the decision maker has some information about C, described by a certain σ-algebra $\mathcal{C}$, and that, on the basis of this information, he is able to rank all the bets in

$$F^+(C) = \{f : C \to \mathbb{R}_+ \mid f \text{ bounded and } \mathcal{C}\text{-measurable}\}$$

Let us denote by $\succsim_C$ the decision maker's ranking of the bets in $F^+(C)$. Now, our decision maker wants to use this knowledge, and possibly the knowledge that Siberia and California are at a different latitude, longitude, that they are both on the Pacific, etc., to solve the problem he faces, namely that of ranking the bets in $\mathcal{F}$.

To begin, let us define Σ as the coarsest σ-algebra which makes all the bets in $\mathcal{F}$ measurable, and (to simplify the exposition) set

$$\mathcal{F} = F^+(S) = \{f : S \to \mathbb{R}_+ \mid f \text{ bounded and } \Sigma\text{-measurable}\}$$

Next, we want to formalize the idea that the problem of ordering $F^+(S)$ can be "translated" into the problem of ordering $F^+(C)$, which has already been solved by the decision maker with the ordering $\succsim_C$. The task boils down to identifying two requirements. First, it must be the case that each bet $f_i \in F^+(S)$ can be identified to a bet $\phi_i \in F^+(C)$. In other words, there must exist a mapping $\nu : F^+(S) \to F^+(C)$. Second, it must be the case that, modulo the renaming of the bets produced by the mapping ν, the objects $F^+(S)$ and $\nu(F^+(S)) \subset F^+(C)$ are the same. This means that if there exists a certain relation between any two bets $f, g \in F^+(S)$, then this relation must be preserved once the mapping ν is applied. That is, it must be the case that the mapping ν preserves whatever structure is associated with $F^+(S)$. In our context, $F^+(S)$ is an affine space of measurable mappings. Hence, the requirement that ν be structure-preserving translates into the demand that ν displays the following properties

(i) affinity: $\nu(\alpha f + \beta g) = \alpha\nu(f) + \beta\nu(g), \qquad \alpha, \beta \in \mathbb{R}_+$

(ii) normality: $f_n \nearrow f \Longrightarrow \nu(f_n) \nearrow \nu(f), \qquad n \in \mathbb{N}.$

That is, if a bet $h \in F^+(S)$ is a combination of two other bets, $f, g \in F^+(S)$, then we want $\nu(h)$ to be an analogous combination of the bets $\nu(f)$ and $\nu(g)$; and if a collection of bets $\{f_n\}$ in $F^+(S)$ approximates a bet $f \in F^+(S)$, then we want the collection $\{\nu(f_n)\}$ to approximate the bet $\nu(f)$. Mappings like ν are known as kernels or generalized transition probabilities (see Appendix A.1 for more about kernels).

Once $F^+(S)$ and $F^+(C)$ are made to correspond to each other by means of a kernel $\nu : F^+(S) \to F^+(C)$, then the problem of ordering $F^+(S)$ can be solved by setting

$$f \succsim g \qquad iff \qquad \nu(f) \succsim_C \nu(g)$$

We summarize the content of this discussion in the following definition.[3]

Definition 1 *Let (S, Σ) and $(C, \mathcal{C})$ be two measurable spaces. Let d be a decision maker who displays choices $\succcurlyeq_S$ and $\succcurlyeq_C$ on $F^+(S)$ and $F^+(C)$, respectively.*

[3] Our definition is in the spirit of [9]. The latter has been criticized in [3]. See also [23] and [10] for more on the debate.

We say that the problem $(F^+(C), \succsim_C)$ is a model for ranking the bets in $F^+(S)$ if there exists a kernel $\nu : F^+(S) \to F^+(C)$ such that

$$f \succcurlyeq_S g \qquad iff \qquad \nu(f) \succcurlyeq_C \nu(g)$$

We say that the problems $(F^+(C), \succsim_C)$ and $(F^+(S), \succsim_S)$ are analogous if the kernel is a bijection.

The definition makes it clear that the concepts of *model* and *analogy* pertain to a decision maker. In other words, one can exhibit two different decision makers for whom the problems on S and C are analogous but the analogy is realized by different kernels as well as a third decision maker for whom there exist no analogy between the two problems.

3 Kernels and models

As a general matter, the possibility of solving a problem, say ordering $F^+(S)$, by analogy with another problem, say $(F^+(C), \succsim_C)$, corresponds to the existence of an affine mapping $\nu : F^+(S) \to \mathbb{R}_+^C$ with the property that $range(\nu) \subseteq F^+(C)$. As the reader has probably already realized, the existence of such a mapping depends crucially on the relation between the two σ-algebras $\mathcal{C}$ and Σ or, equivalently, between $\mathcal{C}$ and the set of bets $\mathcal{F}$ offered to the decision maker.

It is important that the reader keep in mind that the issue, far from being merely technical, is a very substantial one. The σ-algebra $\mathcal{C}$ on C describes (by assumption) the decision maker's understanding of C. This is the same as saying that the decision maker's knowledge is described by $F^+(C)$. Now, suppose that the decision maker attempts to translate his problem into $(F^+(C), \succsim_C)$, and that this produces a certain (non-constant) affine mapping $\nu : F^+(S) \to \mathbb{R}_+^C$. Suppose further that $\mathcal{C}$ is so coarse that this mapping does not satisfy the condition $range(\tilde{\nu}) \subset F^+(C)$. In a manner of speaking, what happens is that $F^+(C)$ is not big enough to accommodate all the functions in $F^+(S)$. This situation has a very transparent meaning. Knowledge of $(F^+(C), \succsim_C)$ is not enough to solve the problem at hand. Equivalently, $(F^+(C), \succsim_C)$ cannot be a model for ordering the bets in $F^+(S)$.

To illustrate the point further, consider the following example. Let $S = [0, 1]$ be equipped with the Lebesgue σ-algebra, and let C be a two-point set, $C = \{c_1, c_2\}$. A mapping $c_i \longmapsto \mu_i \in \Delta(S)$ defines a mapping (see Appendix A.1) $\kappa : F^+([0, 1]) \to \mathbb{R}_+^C$ in a way that each bet f on $[0, 1]$ is associated to the function $\phi \in \mathbb{R}_+^C$ which is the two-coordinate vector $\phi = (\int_{[0,1]} f d\mu_1, \int_{[0,1]} f d\mu_2)$. Clearly, κ is affine. Now, suppose that C is equipped with the trivial algebra $\{\emptyset, C\}$. Then, while the set of real-valued mappings on C is isomorphic (in a set-theoretic sense) to the plane, the set of measurable functions on C consists of the constants only. If $\mu_1 \neq \mu_2$, then it is an easy matter to check that uncountably many measurable functions on $[0, 1]$ are associated to vectors whose first coordinate differs from the second, that is to non-constant functions on C. Since these are not measurable, we see that $range(\kappa) \nsubseteq F^+(C)$.

What drives the example is that while C is a two-point set it behaves, from the viewpoint of the measurable properties, as a one-point space. This property is only seemingly artificial. In fact, as shown in [1], many natural information structures produce spaces of this sort.

4 Ellsberg's three-color urn experiment

Here we present two instances in which an analogy fails to exists. In the first, there is no affine mapping capable of realizing an analogy satisfying certain conditions; in the second, any affine mapping fails the range condition discussed in the previous section. In both cases, the underlying decision problem is Ellsberg's three-color urn experiment [8]. Recall that in the three-color urn experiment, a decision maker is asked to rank bets which pay a certain amount of money depending on the color of a ball which is drawn from an urn. He is told that the urn contains 90 balls, of which 30 are red (R) while the remaining are either blue (B) or green (G). Suppose also that the decision maker is told that the blue balls are either 28 or 32.

As a first example, consider the following story. Suppose that an alien (clearly, out of luck) happens to land on a decision theorist's backyard. The decision theorist, enticed by the rather unique opportunity to test a subject uncontamined by human preconceptions, locks the alien in a room, and performs a first experiment. This goes as follows. A first urn (labeled C) is put on a table, and the alien is told that C contains 30 balls. Next, the decision theorist tells the alien that he has two more urns, one with 28 balls and one with 32. He puts one of those on the table, let us call it U, and asks the alien to point at one urn between C and U. If the urn the alien points at is the one containing the highest number of balls, then the alien will be rewarded. Suppose that, for whatever reason, the alien points at C, that is C$\succ$U. This being done, the decision theorist moves on, and performs Ellsberg's experiment with the alien. The latter, called upon ranking the bets, realizes that the problem of comparing many bets is of the same type as the problem of comparing C and U, the one that he solved before. Driven by this observation, he would associate betting on R with betting on C, betting on B or G with betting on U, betting on $B+G$ with betting on (possibly a multiple of) C, betting on $R+G$ with betting on (possibly a multiple of) U, etc. Here, the requirement that the mapping be affine is clearly important: in particular, it expresses the fact that there is a fixed number of balls in Ellsberg's urn (a fact which we assume is understood by the alien). However, it is an easy matter to check that the alien's mapping cannot be affine. That is, the problem of ranking C and U does not provide a solution for Ellsberg's experiment.

As a second example, suppose that the alien starts out with Ellsberg's experiment. In a rather sophisticated fashion, he realizes that ranking Ellsberg's bets is equivalent to betting on the realization of one of two measure spaces, (S, Σ, μ_1) and (S, Σ, μ_2), where (S, Σ) is the set $\{R, B, G\}$ equipped with the maximal algebra, $\mu_1 = (30, 28, 32)$ and $\mu_2 = (30, 32, 28)$. This is the same as saying that he is considering as a candidate model the set $M = \{\mu_1, \mu_2\}$. However, not knowing much about humans' psychology, the alien might be incapable of assesing which process might have led the decision theorist to select (S, Σ, μ_1) rather than (S, Σ, μ_2) or vice versa. This translates into the fact that the alien's algebra on M is the trivial one. In such a situation, just like in the example of the previous section, any affine mapping $F^+(S) \to \mathbb{R}_+^C$ fails the range condition. In fact, the only acts corresponding to measurable mappings are those in the linear span of "betting on R", "betting on $B + G$" and "betting on S ": the set $F^+(M)$ is not big enough to accommodate all the acts that the alien has to rank. Incidentally, this situation lends itself to an interesting interpretation. Acts corresponding to measurable mappings on M are precisely those acts that are measurable with respect to the decision maker's information. We may call these acts *subjectively measurable*. This class restricted to indicator functions gives us a class of *subjectively measurable events*: on the basis of his information, the decision maker is able to assess the likelihood of these events. In our example, the class consists of the events $\{\varnothing, R, B \cup G, S\}$.[4] We will come back to this point in the concluding remarks.

5 More models

Here, we want to allow for the possibility that our decision maker uses more than one model to solve the problem at hand, that of ordering the bets in $F^+(S)$. For instance, when it comes to forecasting the weather in Siberia, the decision maker might use, not only his knowledge of the weather in California, but also his knowledge of the weather in Nevada as well as that of any relation between the weather in California and Nevada, and so forth. A bit more formally, let M be the set of models for our decision maker. Each model $m \in M$ is a pair $(F^+(C_m), \succsim_{C_m})$, where $(C_m, \mathcal{C}_m)$ is some measurable space, and there is given a kernel $\kappa_m : F^+(S) \to F^+(C_m)$. Each model produces, by means of the associated kernel, an ordering of the bets in $F^+(S)$, and, generally speaking, different models produce different orderings. Then, the decision maker uses this collection of orderings to come up, if possible, with a solution for his problem. Just like before, the decision maker's understanding of M is described by two ingredients: (a) the set of bets on M that the he knows how to order; and (b) the way he orders them. The first is described by a σ-field, $\mathcal{M}$, of subsets of M or, equivalently, by the set $F^+(M)$ of nonnegative, bounded $\mathcal{M}$-measurable functions. The second, by a binary relation, $\succsim_M$, on $F^+(M)$. Intuitively, one might think of the latter as representing the decision maker's assessments of which model is more likely to apply. To simplify the exposition, let us assume that, for each $m \in M$, the ordering $\succsim_{C_m}$ is represented by a functional j_m :

[4] A similar conclusion can be reached in the first example. There the decision maker sets up a mapping taking R into C and B and G into U. Given the maximal algebra on {C,U}, this mapping produces on S exactly the algebra described in the text. In general, however, the two procedures are not equivalent as the class of events obtained with the second procedure is only a λ-system.

$F^+(C_m) \to \mathbb{R}_+$. Such an assumption could be easily dispensed with, but at the cost of introducing a fairly cumbersome notation. Since here we are concerned more with the basic ideas rather than with formalities, the simplification is, in fact, harmless.

Following the idea outlined above, each function $f \in F^+(S)$ is now associated to the function $\phi_f \in \mathbb{R}_+^M$ which at point $m \in M$ takes the value

$$\phi_f(m) = j_m(f)$$

In a sense, the function ϕ_f is a description of the bet f when all models are taken into account. Let $\kappa : F^+(S) \to \mathbb{R}_+^M$ be the function defined by $f \longmapsto \phi_f$. At this point, just like in Section 3, we have two possibilities. Either $range(\kappa) \subset F^+(M)$ or the inclusion does not hold. In the first case, the decision maker solves his problem by setting

$$f \succsim g \qquad iff \qquad \kappa(f) \succsim_M \kappa(g)$$

In the second, he concludes that his understanding of (his information about) M is not enough to solve the problem at hand, and must rely on other considerations.[5]

On the other hand, if it happens that $range(\kappa) \subset F^+(M)$ then the problem of ranking the bets in $F^+(S)$ splits into two parts. First, the mapping κ takes f into $\kappa(f)$, then $\kappa(f)$ is ranked by means of $\succsim_M$. If, to fix ideas, we assume that $\succsim_M$ be itself represented by a functional $V : F^+(M) \to \mathbb{R}_+$, then we can summarize the discussion by means of the following diagram

$$\begin{tikzcd} F^+(S) \arrow[r, "\kappa"] \arrow[dr, "I"'] & F^+(M) \arrow[d, "V"] \\ & \mathbb{R}_+ \end{tikzcd}$$

that is

$$f \succsim g \qquad iff \qquad I(f) \geq I(g)$$

where $I : F^+(S) \to \mathbb{R}_+$ is defined by $I = V \circ \kappa$.

We conclude the section with a couple of remarks about the process of decision making that emerges from our construction. First, we suggest another possible interpretation of this process, which exploits the fact (noted in Section 1) that Case-Based Reasoning is a special case of Reasoning by Analogy. According to this point of view, each element in $m \in M$ can be thought of as a case or, more generally, as a collection of cases on the same domain. Then, the associated ordering on $F^+(S)$ can be thought of as the ordering suggested by the "experience" m and by the mapping κ_m. The set M is the collection of all such experiences and the field $\mathcal{M}$ describes the decision maker's view of

[5] Roughly, it means that the decision maker has to rely on considerations which are "non-measurable" with respect to his information about M.

how all these experiences (and the associated orderings on $F^+(S)$) fit together. We would like to emphasize, however, that in general M may contain, not only cases actually experienced by the decision maker, but also hypothetical cases and even cases the decision maker has never thought of. As a matter of fact, the field $\mathcal{M}$ and the functional V determine which cases in M affect the decision maker's ordering of the acts. For instance, suppose that $A \subset M$ is a subset of cases that the decision maker has never thought of and that, as such, are not going to affect his behavior. If, for instance, V is defined by means of a probability measure P on $\mathcal{M}$, then the condition $P(A) = 0$ would convey that such cases play no role.[6]

6 Choice-based foundations for models and analogies

The whole discussion above was heuristic. Some readers might find our definitions of model and analogy reasonable, others might be sceptical about them. The same can be said about the model of decision making which we outlined at the end of the previous section. It is a fact, nonetheless, that such a model admits a rigorous, choice-based, foundation.

Let $\mathcal{F}_0$ denote the set of simple Σ-measurable acts and $\mathcal{F}_c$ that of constant acts. Let $\succsim$ be a preference relation on $\mathcal{F}_0$ satisfying the following axioms.

A1 $\succsim$ is complete and transitive.

A2 (C-independence) For all $f, g \in \mathcal{F}_0$ and $h \in \mathcal{F}_c$ and for all $\alpha \in (0,1)$

$$f \succ g \qquad \Longleftrightarrow \qquad \alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h$$

A3 (Archimedean property) For all $f, g, h \in \mathcal{F}_0$, if $f \succ g$ and $g \succ h$ then $\exists \alpha, \beta \in (0,1)$ such that $\alpha f + (1-\alpha)h \succ g$ and $g \succ \beta f + (1-\beta)h$.

A4 (Monotonicity) For all $f, g \in \mathcal{F}_0$, $f(s) \succsim g(s) \Longrightarrow f \succsim g$.

A5 (Non-degeneracy) $\exists x, y \in X$ such that $x \succ y$.

Ghirardato, Maccheroni and Marinacci [11] have shown that such preferences are completely described by the functional $I : F^+(S) \to \mathbb{R}$ defined by

$$I(u \circ f) = \alpha(f) \min_{P \in \mathcal{C}} \int_S u \circ f dP + (1 - \alpha(f)) \max_{P \in \mathcal{C}} \int_S u \circ f dP$$

where $\mathcal{C}$ is a weak*-compact set of probability measures, $\alpha : F^+(S) \to [0,1]$ and u is a utility function on the prize space. It is readily seen that as special cases, one obtains α-maxmin expected utility ($\forall \alpha \in [0,1]$), Choquet expected utility and Subjective expected utility. With less emphasis on the form of the functional I and more on the structure that emerges from the theorem, we can reformulate the result as follows.

[6] The converse to this statement is not true. $P(A) = 0$ does not mean that the decision maker did not think of cases in A.

Theorem 2 *Given a preference relation satisfying Axioms 1 to 5 there exist*
(i) a measurable space $(M, \mathcal{M})$
(ii) an affine mapping $\kappa : F^+(S) \to F^+(M)$
(iii) a functional $V : F^+(M) \to \mathbb{R}$
such that

$$f \succcurlyeq_S g \qquad iff \qquad V \circ \kappa(f) \geq V \circ \kappa(g)$$

Moreover, one can take $M = \mathcal{C} \subset \Delta(S)$ *and* κ *is defined by* $f \longmapsto \phi_f$*, where* ϕ_f *is the function that at point* $\mu \in \mathcal{C}$ *takes the value*

$$\phi_f(\mu) = \int_S f d\mu$$

Remark 3 *A utility function on the prize space,* $u : X \to \mathbb{R}_+$*, produces a mapping from the set of acts into* $F^+(S)$ *given by* $f \longmapsto u \circ f$*. In the statement of the theorem, we have identified an act with its image in* $F^+(S)$*. This is harmless and simplifies the notation. We will stick with this convention throughout the paper.*

The theorem says that such a general model of decision making displays the same structure as the model we proposed at end of last section. In light of this observation, those discussions now provide a possible interpretation for the model emerging from Theorem 2. In other words, we can think of elements in M as "models" for the decision maker and of the functional I as consisting of two parts as in the diagram below

$$\begin{array}{ccc} F^+(S) & \xrightarrow{\kappa} & F^+(M, \mathcal{M}) \\ & I \searrow & \downarrow V \\ & & \mathbb{R}_+ \end{array}$$

The first part is the mapping κ, which takes a bet $f \in F^+(S)$ into a function $\phi_f \in F^+(M)$. At model $m \in M$, ϕ_f takes a real value, $\phi_f(m)$, which can be interpreted as the evaluation of bet f corresponding to that model. All such possible evaluations are, then, collected together and lead to a single evaluation of the bet f by means of the functional V.

In [11], Ghirardato, Maccheroni and Marinacci introduced the following relation on $\mathcal{F}$, which they termed unambiguous preference relation.

Definition 4 ([11]) *Let* $f, g \in \mathcal{F}$*.* f *is unambiguously preferred to* g*,* $f \succsim^* g$*, if*

$$\lambda f + (1-\lambda) h \succsim \lambda g + (1-\lambda) h$$

for all $\lambda \in (0,1)$ *and* $h \in F$*.*

As shown in [11] (Sec. B.3), for a preference relation satisfying A1 to A5, the following axiom is equivalent to the property that all the priors in the representation are countably additive.

A6 (Monotone Continuity) For all $x, y, z \in X$ such that $y \succ^* z$, and all sequences of events $\{A_n\}_{n \geq 1} \subseteq \Sigma$ with $A_n \downarrow \emptyset$, there exists $\bar{n} \in \mathbb{N}$ such that $y \succ^* xA_{\bar{n}}z$.

The introduction of Axiom 6 allows us to conclude that the mapping κ in Theorem 2 is a kernel. As noticed, κ is affine. Without Axiom A6, this is all we can say. In particular, we cannot conclude that κ is normal because with finitely additive measure the Dominated Convergence Theorem need not hold. However, with Axiom 6, all the measures in the theorem are countably additive. Hence, since every f is bounded, it is immediate to verify that κ is normal.

7 More general preferences

In this section, we are going to show that theorems of the above type hold for a much wider class of preferences. In Theorem 5 below, we drop axiom A2 (C-independence) and significantly weaken other axioms. Dropping the axiom of C-independence is especially interesting in that it allows us to cover the class of Variational Preferences (Maccheroni, Marinacci and Rustichini [22]) which appears in important applications in Macroeconomics and in Finance (Hansen and Sargent [18]). From a more theoretical perspective, Ghirardato, Maccheroni and Marinacci [12] have shown that C-independence is the crucial property guaranteeing the complete separation between utility and "beliefs" (see [12] with regard to this terminology). Because of this, preferences satisfying A1 to A5 enjoy special properties, and are termed *invariant biseparable preferences*. Then, our Theorem 5 shows that the properties of the previous section hold for all *biseparable preferences* (not necessarily invariant) and beyond that. In Theorem 6, we show that our conclusions extend beyond the class of Archimedean preferences. By this, we mean that the property in axiom A3 fails not only for general acts, but for constant acts too. In such a case, there exists no real-valued utility on the prize space.

To begin, let us introduce two new axioms:

A3' X, the prize space, is a connected, separable topological space. Moreover, for each $x' \in X$, the sets $\{x \in X \mid x \succsim x'\}$ and $\{x \in X \mid x \precsim x'\}$ are closed.

CE (Certainty Equivalents) For all $f \in \mathcal{F}_0$, $\exists x_f \in X : f \sim x_f$.

Theorem 5 *Let $\succsim$ be a preference relation on $\mathcal{F}_0$. If $\succsim$ satisfies A1, A3' and CE, then there exist*
(i) a measurable space $(M, \mathcal{M})$
(ii) an affine mapping $\kappa : F^+(S) \to F^+(M)$
(iii) a functional $V : F^+(M) \to \mathbb{R}$
such that

$$f \succcurlyeq_S g \qquad iff \qquad V \circ \kappa(f) \geq V \circ \kappa(g)$$

Moreover, one can take $M = \Delta(S)$ and κ is defined by $f \longmapsto \phi_f$, where ϕ_f is the function that at point $\mu \in \Delta(S)$ takes the value

$$\phi_f(\mu) = \int_S f d\mu$$

The proof is in Appendix. Just as in the proof of Theorem 2, we use the existence of a utility function on the prize space. Axioms A1 and A3' guarantee the existence of a utility. However, because we have dropped A2, the utility is not necessarily linear.

We also observe that, as a special case, the same statement holds if we reintroduce A4 and replace CE with the continuity assumption S1 in [14], Section 3.3. For, in such a case, Lemma 29 in [14] implies that axiom CE is automatically satisfied. Such preferences are not necessarily biseparable in that they do not have to satisfy the assumption of Binary Comonotonic Independence in [14], Section 3.3.

We now move to the non-Archimedean case. We first introduce the following axiom, which weakens A2 in that it requires that independence hold for elements in the prize space only.

LC For all $x, y, z \in X$ and for all $\alpha \in (0,1)$

$$x \succ y \qquad \Longleftrightarrow \qquad \alpha x + (1-\alpha)z \succ \alpha y + (1-\alpha)z$$

Given axiom LC, the extension of the above results to the non-Archimedean case obtains along similar lines. In fact, axioms A1 and LC guarantee the existence of a utility representing the ordering on the prize space. Since we gave up the Archimedean axiom, such a utility is not valued in the reals but rather in an ordered vector space OV^* (see Hausner [19]). Without loss, we can assume that the latter is a Banach space. Hence, acts can be identified to mappings from S to this Banach space. By axiom CE, we can define an operator from such mappings to OV^*, which represents the decision maker's ordering of the acts. Finally, we define κ and V just like we did in the previous theorems. The only difference is that the integrals of the previous theorems as well as the mapping V are now OV^*-valued (the integrals are, in fact, Bochner integrals). Summarizing, we have

Theorem 6 *Let $\succsim$ be a preference relation on $\mathcal{F}_0$. If $\succsim$ satisfies A1, A4, CE and LC, then there exist*
(o) a utility function $u : X \to OV^*$, where OV^* is a Banach space.
(i) a measurable space $(M, \mathcal{M})$
(ii) an affine mapping $\kappa : F_0(S, OV^*) \to F(M, OV^*)$ [7]
(iii) a mapping $V : F(M, OV^*) \to OV^*$
such that

$$f \succcurlyeq_S g \qquad iff \qquad V \circ \kappa(f) \geq V \circ \kappa(g)$$

[7] $F_0(S, OV^*)$ is the set of simple measurable mappings $S \to OV^*$ and $F(M, OV^*)$ that of measurable mappings $M \to OV^*$.

Moreover, one can take $M = \Delta(S)$ and κ is defined by $f \longmapsto \phi_f$, with ϕ_f being the function that at point $\mu \in \Delta(S)$ takes the value

$$\phi_f(\mu) = \int_S f d\mu$$

where the integral on the RHS is a Bochner integral.

8 Concluding remarks

We have shown that many preferences are compatible with the idea of analogical reasoning presented here. The main weakness of our result is obvious: neither we proved the uniqueness of the axiomatization associated to our model, nor any uniqueness result can be obtained on the basis of the data we used. In other words, while the preferences we have studied are certainly compatible with the resoning process we have proposed, they might be compatible with other types of reasoning as well, and nothing in this paper would tell how to screen among different possibilities. We believe that much tighter results might be obtained by explicitly modeling a dynamic process where, very much in the spirit of Gilboa and Schmeidler [17], one keeps track of the problems solved by the decision maker, of the outcomes realized in those and where the idea of analogical reasoning is coupled with a learning paradigm compatible with it. This goes, however, beyond the scope of this paper. Yet, it stays that, in a large variety of cases, we can think of individuals as making their choices by drawing analogies in the way described here. Probably, the strength of our definition lies herein. If the reader agrees that analogy is a pervasive process in human life (see Section 1), then he would probably agree that any definition that does not accomplish this is suspicious, to say the least. From this perspective, our results say that our definition is a suitable contender for a "right" definition, and might serve as basis for further work (for instance, a dynamic extension a la' [17]). After all, the study of reasoning processess underlying decision making is still a largely unexplored territory.

8.1 Analogy, ambiguity and unforeseen contingencies

In Section 4, we saw examples where an analogy fails to exist: a decision maker was trying to establish an analogy between the problem of ranking $F^+(S)$ and another problem (or set of problems) $(F^+(C), \succsim_C)$ but the mapping $\nu : F^+(S) \to \mathbb{R}_+^C$ failed the range condition discussed in Section 3. In this circumstances, we are led to distinguish between acts in $\nu^{-1}(F^+(C))$ and acts in $F^+(S) \backslash \nu^{-1}(F^+(C))$: while acts in $\nu^{-1}(F^+(C))$ – which we call subjectively measurable – can be evaluated on the basis of the information encoded in $(F^+(C), \succsim_C)$, acts in $F^+(S) \backslash \nu^{-1}(F^+(C))$ cannot. As such, this distinction is reminiscent of that between unambiguous and ambiguous acts, which has been studied by several scholars. In [1], it is shown that the class of subjectively measurable events (the class of indicator functions in $\nu^{-1}(F^+(C))$ is always a λ-system and always contains the class of unambiguous events in the sense of [11] and [24]. We refer the reader to [1] for more details.

Situations where an analogy fails to to exist lend themselves to a somewhat "dual" interpretation linking the concept of analogy to that of unforeseen contingencies [5]. The idea is that, in actual circumstances, individuals might erroneously establish an analogy between the problem at hand and another problem they solved before, but by doing so they would "misrepresent" the problem they face. For instance, let us consider an individual who is in the process of negotiating a contract with another party, and let us suppose that, *from the viewpoint of a perfectly informed outside observer*, the individual's problem is equivalent to that of ranking all acts $F^+(S)$ on a certain measurable space (S, Σ). In actual situations, however, there is no guarantee that the individual would recognize such equivalence. In fact, one of the most arduous tasks one typically faces is that of coming up with a correct description of the problem at hand; that is, to understand what are the alternatives available, what are the relevant contingencies, etc.. It is then possibile – maybe even common – that our individual might erroneously think that the problem he faces is analogous to some other problem, say $(F^+(C), \succsim_C)$, that he solved before. In this case, he would recognize only alternatives that are in the set $\nu^{-1}(F^+(C))$, and just would not "think" of those in $F^+(S)\backslash\nu^{-1}(F^+(C))$ (as an extreme case, the individual might conclude that he has only one choice while other alternatives are, in fact, available). The underlying class of relevant events would also be constructed following the same procedure, and events (=indicator functions) in $\Sigma\backslash\nu^{-1}(F^+(C))$ would describe events that he just did not think of.

References

[1] Amarante M. (2005), States, Models and Unitary Equivalence: A reconsideration of Ellsberg's Paradox, Columbia University, mimeo.

[2] Anscombe F.J. and R.J. Aumann (1963), A definition of subjective probability, *Annals of Mathematical Statistics* **34**, 199-205.

[3] Chalmers D.J., R.M. French and D.R. Hofstadter (1992), High-level perception, representation, and analogy: A critique of AI methodology", *Journal of Experimental & Theoretical Artificial Intelligence* **4**, 185-211.

[4] Debreu G. (1954), Representation of a preference ordering by a numerical function, in Decision Process, R.M. Thrall, C.H. Coombs and R.L. Davis eds, John Wiley.

[5] Dekel E., B. Lipman and A. Rustichini (2001), Representing Preferences with a Unique Subjective State Space, *Econometrica* **69**, 891-934.

[6] Diestel J. and J. J. Uhl (1977) Vector Measures, AMS Mathematical Surveys No. 15.

[7] Dictionary of Philosophy of Mind, www.artsci.wustl.edu/~philos/MindDict/analogy.html

[8] Ellsberg, D. (1961), "Risk, Ambiguity and the Savage Axioms", *Quarterly Journal of Economics* 75, 643-69.

[9] Falkenheimer B., K.D. Forbus and D. Gentner (1989), The Structure mapping engine: algorithm and examples, *Artificial Intelligence* **41**, 1-63.

[10] Forbus K.D., D. Gentner, A.B. Markman and R.N. Ferguson (1998), Analogy just looks like high level perception: Why a domain-general approach to analogical mapping is right, *Journal of Experimental & Theoretical Artificial Intelligence* **10:2**, 231-57.

[11] Ghirardato P., F. Maccheroni and M. Marinacci (2003), Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory*, forthcoming.

[12] Ghirardato P., F. Maccheroni and M. Marinacci (2003), Certainty Independence and the Separation of Utility and Beliefs, *Journal of Economic Theory*, forthcoming.

[13] Ghirardato P., F. Maccheroni, M. Marinacci and M. Siniscalchi (2003), Subjective Foundations for Objective Randomization: A New Spin on Roulette Wheels, *Econometrica* **71**, 1897-1908.

[14] Ghirardato P. and M. Marinacci (2001), Risk, ambiguity and the separation of utility and beliefs, *Mathematics of Operation Research* **26**, 864-90.

[15] Gilboa I. and D. Schmeidler (1989), Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics* **18**, 141-53.

[16] Gilboa I. and D. Schmeidler (2001), A Theory of Case-Based Decisions, Cambridge University Press.

[17] Gilboa I. and D. Schmeidler (2003), Inductive Inference: An Axiomatic Approach, *Econometrica* **71**, 1-26.

[18] Hansen L. and T. Sargent (2000), Wanting robustness in macroeconomics, mimeo.

[19] Hausner M. (1954), Multidimensional Utilities, in Decision Process, R.M. Thrall, C.H. Coombs and R.L. Davis eds, John Wiley.

[20] Hausner M. and J.G. Wendel (1952), Ordered vector spaces, *Proceedings of the AMS* **3**, 977-82.

[21] Kastler D. (1982), On A. Connes' Noncommutative Integration Theory, *Communications in Mathematical Physics*, **85**.

[22] F. Maccheroni, M. Marinacci and A. Rustichini (2004), Variational representation of preferences under ambiguity, ICER WP 5/04

[23] Morrison C.T. and E. Dietrich (1995), Structure-mapping vs. high-level perception: the mistaken flight over the explanation of Analogy", *Proc. 17th Annual Conference of Cognitive Science Society*, 678-82.

[24] Nehring K. (2001), Ambiguity in the context of probabilistic beliefs, UC Davis, mimeo.

APPENDIX

A.1 Kernels

The concept of kernel is the building block of our definition of reasoning by analogy. In fact, kernels play an important role throughout this work. Thus, it is appropriate to gather here a few basic facts about them. For more, the reader should consult [21].

Let $(Y, \mathcal{Y})$ and $(Y', \mathcal{Y}')$ be two measurable spaces. A kernel from Y to Y' is a mapping $\nu : F^+(Y) \to F^+(Y')$ which is

(i) affine: $\nu(\alpha f + \beta g) = \alpha\nu(f) + \beta\nu(g), \qquad \alpha, \beta \in \mathbb{R}_+$

(ii) normal: $f_n \nearrow f \Longrightarrow \nu(f_n) \nearrow \nu(f), \qquad n \in \mathbb{N}$

In other words, a kernel is a representation (that is, a mapping which is structure-preserving) of the positive measurable functions on Y into the positive measurable functions on Y'. Any kernel can be equivalently described by means of a mapping, $Y' \to \Delta^+(Y)$, from Y' to the set, $\Delta(Y)$, of positive measures on Y. In fact, one associates the element $y' \in Y'$ to the measure $\nu^{y'} \in \Delta(Y)$, which is defined by the equation

$$\nu^{y'}(A) = (\nu(\chi_A))(y'), \qquad for \quad every \quad A \in \mathcal{Y}$$

and the mapping

$$Y' \to \Delta^+(Y) \qquad defined \quad by \qquad y' \longmapsto \nu^{y'} \tag{1}$$

is *measurable* in the sense that for every $A \in \mathcal{Y}$, the function on Y' which takes the value $\nu^{y'}(A)$ at the point $y' \in Y'$ is an element of $F^+(Y')$.

The converse of this is not necessarily true. While any mapping $Y' \to \Delta(Y)$ defines a mapping $\tilde{\nu} : F^+(Y) \to \mathbb{R}^{Y'}$, it is not guaranteed that $range(\tilde{\nu}) \subset F^+(Y')$. That is, $\tilde{\nu}$ might fail to be measurable in the above sense. For instance, given $Y' \to \Delta(Y)$, we can define $\tilde{\nu}$ by

$$\tilde{\nu} : f \longmapsto \tilde{\nu}(f)$$

where $\tilde{\nu}(f)$ is the function in $\mathbb{R}^{Y'}$ which at point $y' \in Y'$ takes the value

$$\tilde{\nu}(f)(y') = \int_Y f d\nu^{y'} \tag{2}$$

and it is clear that whether or not $\tilde{\nu}(f)$ is a measurable function on Y', i.e. is an element of $F^+(Y')$, depends on the σ-algebra $\mathcal{Y}'$ that we have on Y'. On the other hand, it is immediate to check that if $range(\tilde{\nu}) \subset F^+(Y')$, then $\tilde{\nu}$ is affine and normal and, therefore, it is a kernel, and all kernels are defined essentially in this way (see equation (2)).

A.2 Omitted proofs

Recall that S is the domain of the acts, Σ is a fixed field of events in S and X is the prize space. We assume that X is a convex set. As it is well-known, such an assumption can be justified by thinking of X as the set of lotteries on some given set of outcomes as in Anscombe and Aumann [2]. Alternatively, the assumption can be justified on the basis of the axiomatization of preferences given in Ghirardato, Maccheroni, Marinacci and Siniscalchi [13].

Let $\mathcal{F}_0$ be the set of simple Σ-measurable acts (acts that take only finitely many values in X) and let $\mathcal{F}_c$ be the set of constant acts. Let $\succsim$ be a binary relation on $\mathcal{F}_0$, and let $\succ$ denote its asymmetric part.

Elements in the prize space are identified to the set of constant acts. Axioms 1 to 4 in the text imply the existence of a linear utility on the prize space. As noticed (remark 3), this allows to identify acts with the set $B_0(S, W)$ of bounded Σ-measurable simple functions on S which take values in $W = range(u)$. Gilboa and Schmeidler [15] have shown that preferences satisfying A1 to A5 are represented by a functional $I : B_0(S, W) \to \mathbb{R}$ which is C-independent, positively homogeneous (which allows to extend by homogeneity I to the whole B_0 – the set of bounded Σ-measurable simple functions on S), monotone and supnorm continuous. Finally, the latter property allows to extend I to the set of bounded Σ-measurable functions on S as B_0 is norm dense in the latter set.

Proof of Theorem 2. In [11], it was shown that I takes the form

$$I(u \circ f) = \alpha(f) \min_{P \in \mathcal{C}} \int_S u \circ f dP + (1 - \alpha(f)) \max_{P \in \mathcal{C}} \int_S u \circ f dP$$

for $\mathcal{C}$ a weak*-compact set of probability measures, $\alpha : F^+(S) \to [0, 1]$ and u a utility function on the prize space.

Set $M = \mathcal{C} \subset \Delta(S)$, and let $\mathcal{M}$ be the Borel tribe generated by the weak*-topology on $\Delta(S)$. Then, let κ be defined by $f \longmapsto \kappa(f)$ where

$$\kappa(f)(\mu) = \int f d\mu \qquad , \qquad \mu \in \Delta(S)$$

Notice that $\mathcal{C}$ is measurable in $\Delta(S)$ because it is weak*-closed.

In order to prove the theorem, we need to show two things:

(i) κ is measurable in the sense of A.5.1, i.e. $range(\kappa) \subset F^+(\mathcal{C})$;

(ii) the function $\alpha : F^+(S) \to [0, 1]$ is compatible with the nucleous of equivalence of κ, that is if $f, g \in F^+(S)$ are such that $\kappa(f) = \kappa(g)$, then $\alpha(f) = \alpha(g)$. This will allow us to define α on $F^+(\mathcal{C})$ rather than on $F^+(S)$.

To prove (i) observe that the function $\kappa(f) : \Delta(S) \to R_+$ is trivially continuous for the weak*-topology on $\Delta(S)$. Hence, it is measurable for the Borel tribe generated by that topology.

(ii) was already observed by [11]. We provide here a slightly different proof. Let $f, g \in F^+(S)$ be such that $\kappa(f) = \kappa(g)$. In the terminology of [11], this implies that f is unambiguously indifferent to g. Since unambiguous preference is a subrelation of the decision maker's preference relation over acts, this implies that f is indifferent to g, that is $I(f) = I(g)$. Since $\forall f \in F^+(S)$, $\alpha(f)$ is uniquely defined by

$$\alpha(f) = \frac{I(f) - \max\limits_{\mu \in \Delta(S)} \kappa(f)(\mu)}{\min\limits_{\mu \in \Delta(S)} \kappa(f)(\mu) - \max\limits_{\mu \in \Delta(S)} \kappa(f)(\mu)}$$

we have $\alpha(f) = \alpha(g)$. ■

Proof of Theorem 5. Axioms A1 and A3' imply the existence of a utility function (not necessarily linear) on the prize space ([4]). By axiom CE, for each $f \in \mathcal{F}_0$there exists $x_f \in X$ such that $f \sim x_f$. Define $J : F_0 \to \mathbb{R}$ by $J(f) = u(x_f)$. Clearly, J represents $\succsim$. Set $W = range(u)$. Let $B_0(\Sigma, W)$ denote the set of bounded, Σ-measurable simple functions with range in W. The utility function $u : X \to R$ defines an operator $T_u : F_0 \to B_0(\Sigma, W)$ by $T_u(f) = u \circ f$. Define $I : B_0(\Sigma, W) \to R$ as the unique operator that makes the diagram below commute

$$\begin{tikzcd} F_0 \arrow[r, "T_u"] \arrow[dr, "J"'] & B_0(\Sigma, W) \arrow[d, "I"] \\ & \mathbb{R} \end{tikzcd}$$

$B_0(\Sigma, W)$ is a subset of B_0, the set of bounded, Σ-measurable simple functions. B_0 equipped with the supnorm is Banach space. Hence, it has *sufficiently many* continuous linear functionals. That is, if $a, b \in B_0$, $a \neq b$, there exists a continuous linear functionals L on B_0 such that $L(a) \neq L(b)$. By Riesz representation theorem, a continuous linear functional on B_0 has the form $\int a d\mu$, with μ a finitely additive measure on Σ. Hence, the mapping $\kappa : B_0(\Sigma, W) \to F(\Delta(S))$ defined as in the statement of the proposition is one-to-one. As observed in the proof of the previous theorem, $\kappa(a)$ is a measurable function on $(M = \Delta(S), \mathcal{M})$ with $\mathcal{M}$ being the Borel tribe generated by the weak*-topology. Finally, define V as the unique functional that makes the following diagram commute

$$\begin{tikzcd} B_0(\Sigma, W) \arrow[r, "\kappa"] \arrow[dr, "I"'] & F^+(M, \mathcal{M}) \arrow[d, "V"] \\ & \mathbb{R} \end{tikzcd}$$

■

Proof of Theorem 6. By A1, A4 and LC, the set X along with the order and the mixture operations is a (non-Archimedean) utility space in the sense on Hausner [19]. Hausner [19] and Hausner and Wendel [20] have shown that there exists an order-preserving embedding $\tilde{u}$ (a utility, not real-valued) of X into an ordered vector space OV. Denote by $\geq$ the order on OV.

Define a norm $|\cdot|$ on OV, and denote by OV^* its completion. That is, OV^* is a Banach space. By embedding OV into OV^*, each act $f \in \mathcal{F}_0$ is associated to the mapping $\tilde{u} \circ f$ from S to the Banach space OV^*. For each $f \in \mathcal{F}_0$, $\tilde{u} \circ f$ is measurable and, being a simple function, is strongly measurable. Let $F_0(S, OV^*)$ be the set of all measurable simple mappings $S \to OV^*$.

Endow $F_0(S, OV^*)$ with the norm defined by

$$\|f\| = \{\sup |f(s)| : s \in S\} \qquad , \qquad f \in F_0(S, OV^*)$$

Since OV^* is a Banach space so is $F_0(S, OV^*)$.

By CE (proceeding just like in the previous two proofs), there exists a mapping $J : \mathcal{F}_0 \to OV^*$ such that $f \succsim g$ iff $J(f) \geq J(g)$.

For $\mu \in \Delta(S)$ and $a \in F_0(S, OV^*)$, denote by $\int a d\mu$ the Bochner integral (see for instance [6]).

On $M = \Delta(S)$, we define a topology τ as the coarsest topology such that for any bounded strongly measurable function $b : S \to OV^*$

$$\mu_n \to \mu \qquad \Longrightarrow \qquad \int b d\mu_n \to \int b d\mu$$

Let $\mathcal{B}$ be the Borel tribe generated by τ and let $F(\Delta(S), OV^*)$ be the space of measurable mappings $M \to OV^*$ (these are the mappings that are measurable with respect to $\mathcal{B}$ and the Borel tribe generated by the norm topology on OV^*).

Define $\kappa : f \longmapsto \phi_f$ as in the statement of the theorem. The mapping is clearly injective [If $f, g \in F_0(S, OV^*)$ and $f \neq g$ then (since they are both simple) there exists $A \in \Sigma$ such that $f(s) = x \in OV^*$ and $g(s) = y \in OV^*$for any $s \in A$ and $x \neq y$. Pick $\mu \in \Delta(S)$ so that $\mu(A) = 1$. Then, $\int f d\mu \neq \int g d\mu$]. Moreover, for each $f \in F_0(S, OV^*)$ the mapping ϕ_f is continuous for the topology τ and, hence, measurable. Hence, $range(\kappa) \subset F(\Delta(S), OV^*)$. ■